THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 37, pp. 32464–32474, September 16, 2011
Printed in the U.S.A.

Membrane Localization of Small Proteins in *Escherichia coli*[*] [S]

Received for publication, March 31, 2011, and in revised form, July 20, 2011 Published, JBC Papers in Press, July 21, 2011, DOI 10.1074/jbc.M111.245696

Fanette Fontaine[1,2], **Ryan T. Fuchs**[1,3], **and Gisela Storz**[4]

From the Cell Biology and Metabolism Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892

Escherichia coli **synthesize over 60 poorly understood small proteins of less than 50 amino acids. A striking feature of these proteins is that 65% contain a predicted α-helical transmembrane (TM) domain. This prompted us to examine the localization, topology, and membrane insertion of the small proteins. Biochemical fractionation showed that, consistent with the predicted TM helix, the small proteins generally are most abundant in the inner membrane fraction. Examples of both N_{in}-C_{out} and N_{out}-C_{in} orientations were found in assays of topology-reporter fusions to representative small TM proteins. Interestingly, however, three of nine tested proteins display dual topology. Positive residues close to the transmembrane domains are conserved, and mutational analysis of one small protein, YohP, showed that the positive inside rule applies for single transmembrane domain proteins as has been observed for larger proteins. Finally, fractionation analysis of small protein localization in strains depleted of the Sec or YidC membrane insertion pathways uncovered differential requirements. Some small proteins appear to be affected by both Sec and YidC depletion, others showed more dependence on one or the other insertion pathway, whereas one protein was not affected by depletion of either Sec or YidC. Thus, despite their diminutive size, small proteins display considerable diversity in topology, biochemical features, and insertion pathways.**

Small proteins of less than 50 amino acids have generally been overlooked in all organisms. It is difficult to detect these small proteins using standard biochemical techniques (1). Furthermore, the corresponding genes are frequently ignored in genome annotation and in genetic screens. As a result, the fact that at least 60 of these small proteins are synthesized in *Escherichia coli* had not been appreciated until recently (2). Although the function of most of these small proteins is unknown, many of these proteins are conserved and their expression is regulated, suggesting that they may have important cellular roles.

As increasing numbers of the small proteins are being characterized, it is becoming evident that they can have a wide range of cellular roles (3). For example, *E. coli* MgrB is a 47-amino acid protein that acts as a negative feedback regulator of the PhoQ/PhoP two-component signaling system by interacting with PhoQ in the membrane (4). In *Bacillus subtilis*, the soluble 46-amino acid Sda protein blocks a signaling cascade that leads to sporulation by inhibiting KinA autophosphorylation (5). SpoVM, a 26-amino acid protein that forms an amphipathic helix, associates with the positive curvature of a developing spore in *B. subtilis* and tethers the developing coat to the spore (6). In *E. coli*, the membrane-associated 29-amino acid KdpF and 49-amino acid YbhT proteins stabilize and/or modulate the KdpABC potassium transporter and AcrA-AcrB-TolC efflux pumps, respectively (7).[5]

Intriguingly, most of the small proteins listed above are membrane localized, and almost 65% of the identified small proteins in *E. coli* are predicted to contain an α-helical transmembrane (TM)[6] domain (2). It is likely that correct membrane localization of the proteins is critical to their activities and that information about subcellular localization can give insights into the functions of small proteins. We thus set out to examine whether the small TM domain-containing proteins in *E. coli* are within the inner or outer membrane as well as to determine the orientation of the proteins within the membrane.

We also wanted to identify the pathways used to insert the proteins in the membrane. In *E. coli*, two major membrane insertion systems have been characterized. The system required for most proteins is the Sec translocase (8). This system is comprised of a protein complex, SecYEG, which forms a pore in the inner membrane. Accessory proteins assist Sec substrates in associating with the SecYEG pore and in releasing the substrate from the pore into the inner membrane or moving the substrate through the pore into the periplasm. One of these accessory proteins, YidC, also can act on its own as a separate system of membrane insertion (9). The overall number of substrates solely dependent upon YidC appears to be far fewer than that of Sec substrates, however, several relatively small proteins such as the bacteriophage proteins Pf3 and M13 as well as the *E. coli* protein AtpE, a component of ATP synthase have been reported to require only YidC for membrane insertion (10, 11). Although the features that influence pathway choice remain poorly understood, we predicted that YidC was the most probable insertion pathway for the small proteins characterized here. However, our results with *in vivo* depletion of SecE and

[*] This work was supported, in whole or in part, by the National Institutes of Health Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S5 and Tables S1–S6.

[1] Both authors contributed equally to this work.
[2] Present address: UPR CNRS 9073, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France, Affiliated with Université Paris Diderot - Paris 7, France.
[3] Present address: New England Biolabs, 240 County Rd., Ipswich, MA 01938.
[4] To whom correspondence should be addressed: 18 Library Dr., Bethesda, MD 20892-5430. Tel.: 301-402-0968; Fax: 301-402-0078; E-mail: storz@helix.nih.gov.

[5] E. Hobbs, manuscript in preparation.
[6] The abbreviations used are: TM, transmembrane; SPA, sequential peptide affinity; PhoA, alkaline phosphatase.

YidC show that the situation is more complex; small proteins appear to be inserted into the membrane by a variety of mechanisms.

EXPERIMENTAL PROCEDURES

Plasmids and Strains—All plasmids and strains used in this study are listed under supplemental Tables S1 and S2, respectively. The C-terminal alkaline phosphatase and green fluorescent protein fusions were all generated by amplifying the corresponding control and small protein genes by PCR using oligonucleotides (Integrated DNA Technologies) listed under supplemental Table S3, digesting with XhoI and BamHI and cloning into the corresponding sites of pHA4 and pWALDO, respectively (12, 13).

All strains are derivatives of laboratory stocks of the *E. coli* K12 strain MG1655. Strains with sequential peptide affinity (SPA)- or 3× FLAG tag fusions were made by mini-λ Red recombination (14) as described previously (2) using the oligonucleotides listed under supplemental Table S3. For the *atpE*-SPA fusion, the chromosomal region from 198 nt upstream of the *atpI* start codon to 3 nt downstream of the *atpE* stop codon was amplified and inserted into the *lacZ* gene to make a merodiploid strain. Then, the SPA tag was attached to the copy of *atpE* located in the disrupted *lacZ* gene locus. The *phoA* gene also was deleted using mini-λ Red recombination (14).

To replace the native *yidC* and *secE* promoters with an arabinose inducible promoter, the P_{BAD} cassette from pTM26 (15) was amplified such that the PCR product was flanked by 40 nt of homology to the region of insertion. The PCR products were incorporated into the chromosome of MG1655 mini-λ *tet* ara^+ (GSO497) by mini-λ Red recombination (14) into MG1655 mini-λ *tet* ara^+ (GSO497). For *yidC*, the insertion replaced the region between 66 and 95 nt upstream of the start codon. For *secE*, the insertion replaced the region between 169 and 170 nt upstream of the start codon. The two promoter fusions then were moved into a derivative of MG1655 (GSO468) in which the native *araE* promoter on the chromosome was replaced by the constitutive P_{CP18} promoter (16) to allow for more homogenous expression of *araE*. The presence of the P_{BAD}-*secE* and P_{BAD}-*yidC* cassettes was confirmed by PCR and the loss of ability to grow on LB plates containing 0.2% glucose (supplemental Fig. S1). SPA fusions were moved into the P_{CP18}-*araE* P_{BAD}-*yidC* and P_{CP18}-*araE* P_{BAD}-*secE* strains by P1 transduction (17).

Subcellular Fractionation Using a Sucrose Cushion—Subcellular fractionation based on membrane density was carried out by merging previously published protocols (2, 18). Cell pellets collected for exponentially growing strains expressing SPA-tagged proteins were resuspended in 1 ml of fractionation buffer (100 mM NaCl, 50 mM Tris-HCl, pH 8.0, 20% sucrose), transferred to a microcentrifuge tube, whereupon EDTA (5 mM) and lysozyme (2 mg/ml) were added. Tubes were incubated at 25 °C for 1 h with gentle shaking to allow formation of spheroplasts. The cells were then collected in a microcentrifuge at 20,000 × *g* and resuspended in water to lyse the spheroplasts. The resulting crude lysate was passed through a 30-gauge syringe needle 6 times to homogenize the sample and reduce viscosity. Cell pellets collected for exponentially growing strains expressing 3× FLAG-tagged proteins were resuspended in 4 ml of fractionation buffer and lysed by sonication (Branson Sonifier 450 Branson; duty cycle = 50% intensity, output control = 3) 2 times, 1 min each to obtain the crude lysate.

In both cases, 1 ml of crude lysate was centrifuged at 20,000 × *g* to remove unlysed cells and cellular debris. The supernatant was clarified an additional 3–4 times by repeating the centrifugation. A 500-μl fraction of the clarified lysate was layered on top of a 500-μl sucrose cushion (5 mM EDTA and 1.4 M sucrose). Tubes were centrifuged at 130,000 × *g* for 2 h at 4 °C in a Beckman Optima TLX tabletop centrifuge with a TLA100.3 rotor. After centrifugation, ~425 μl was carefully removed from the top layer (soluble fraction) without disrupting the interface. Then, the interface and remaining liquid was removed (inner membrane fraction). The remaining pelleted material was resuspended in 500 μl of fractionation buffer (pellet fraction). SDS was added to all fractions (final concentration 1%) and the samples were incubated overnight at room temperature. Samples were vortexed to ensure complete solubilization, mixed with loading buffer (1× stacking buffer, 2% SDS, 0.025 mg of bromphenol blue, 52% glycerol), and heated at 95 °C for 5 min prior to loading onto Novex 10–20% Tris glycine gels (Invitrogen). There is some cross-contamination of the fractions because the soluble fraction is not entirely removed after centrifugation and the separation of inner and outer membrane by the cushion is not complete.

Subcellular Fractionation Using Sarcosyl—Subcellular fractionation by differential detergent solubilization was carried out as described (19). Cells were harvested in exponential phase, resuspended in 1× PBS, and lysed by sonication as above. Unlysed cells were removed by a series of three centrifugations at 20,000 × *g* for 5 min. Cell envelopes were separated from the cytosol by centrifugation at 130,000 × *g* for 50 min at 4 °C in the Beckman tabletop ultracentrifuge (TLA100.3 rotor). To separate the outer and inner membranes, the pellet was resuspended in 3 mM EDTA (pH 7.2) and incubated in the presence of 0.5% Sarcosyl for 30 min at room temperature. This mixture was recentrifuged at 130,000 × *g* for 50 min at 4 °C. The supernatant corresponds to the inner membrane. The pellet, which contained outer membranes, was resuspended in 50 mM Tris-HCl (pH 8.0), 10 mM EDTA. Proteins were collected by TCA precipitation, resuspended in loading buffer, and separated as described above.

Immunoblot Assays—Proteins were transferred from denaturing gels to nitrocellulose membranes (Invitrogen), and the membranes were blocked with 3% milk. For SPA- and 3× FLAG-tagged proteins, the membranes were incubated with a 1:1,000 dilution of anti-FLAG M2-AP monoclonal antibody (Sigma), and the protein signals were visualized using Lumi-Phos WB (Pierce). For GFP-tagged proteins, membranes were incubated a 1:10,000 dilution of rabbit anti-GFP antibody, followed by incubation with a 1:10,000 dilution of anti-rabbit horseradish peroxidase. For PhoA-tagged proteins, membranes were incubated with a 1:5,000 dilution of anti-PhoA horseradish peroxidase (Novus Biologicals). To detect the SecY and YidC proteins, membranes were incubated with 1:5,000 dilution of rabbit anti-SecY antibody and rabbit anti-YidC antibody, respectively, followed by incubation with a 1:10,000 dilution of anti-rabbit horseradish peroxidase. All membranes were

exposed to Blue XB film (Kodak), and the protein signals were quantified and analyzed using ImageJ software (National Institutes of Health). For signal quantification, fractions were diluted and care was taken to only quantify bands that were not saturated on the film.

Alkaline Phosphatase Assays—Alkaline phosphatase activity was measured for MG1655 Δ*phoA* transformed with the pHA4 constructs carrying C-terminal PhoA fusions under control of the P_{BAD} promoter. Overnight cultures were diluted 1:100 in 4 ml of LB with 100 μg/ml of ampicillin and incubated at 37 °C for 90 min. The cells were then grown in the presence of 0.2% arabinose for 1 h. An aliquot (1 ml) of cells was harvested and washed with ice cold buffer (10 mM Tris-HCl, pH 8, and 10 mM $MgSO_4$) and resuspended in 1 ml of 1 M Tris-HCl (pH 8). The activity was assayed as described (20). Mean activity values were calculated from at least four independent measurements.

GFP Measurements—GFP activity was measured for BL21(DE3) pLysS cells transformed with the pWALDO constructs carrying C-terminal fusion proteins under control of the T7 promoter. Overnight cultures were diluted 1:50 in 6 ml of LB with 25 μg/ml of chloramphenicol and 50 μg/ml of kanamycin and incubated at 37 °C for 90 min. Cells were grown for another 4 h in the presence of 0.4 mM isopropyl 1-thio-β-D-galactopyranoside to induce T7 RNA polymerase expression. Thereafter, cells were harvested and resuspended in 300 μl of buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 15 mM EDTA and incubated 30 min at room temperature. Aliquots of 100 μl were transferred to ViewPlate 96, White (Packard) 96-well plates, and analyzed for GFP fluorescence emission with a SYNERGY (BioTek) microtiter plate reader (excitation filter 490 nm and emission filter 520 nm). For each sample, the fluorescence activities were normalized by the $A_{600\ nm}$. Mean activity values were obtained from at least three independent measurements.

Proteinase K Assays—Cells expressing the PhoA fusion proteins were grown as for alkaline phosphatase assays, whereas cells expressing SPA fusion proteins were grown to exponential phase. Cells were washed with LB and then resuspended in ice-cold buffer (40% sucrose, 33 mM Tris-HCl, pH 8.0) and incubated with lysozyme (100 μg/ml) and 2 mM EDTA for 15 min on ice. Aliquots of the spheroplast suspension were incubated 1 h on ice without treatment, with proteinase K (0.5 mg/ml), or with proteinase K together with Triton X-100 (2.5%). After the addition of phenylmethylsulfonyl fluoride (2 mM), the PhoA fusion samples were acid-precipitated (trichloroacetic acid, 15% final) and resuspended in 1× loading buffer, whereas 4× loading buffer was directly added to the SPA fusion samples. Proteins were separated on denaturing gels as described above.

Microscopy—Cells expressing the GFP fusion proteins where grown as for the GFP measurements and then spread on a thin layer of agarose between a slide and coverslip and examined using an Axioplan II (Zeiss) microscope. Fluorescence images were acquired with the maximum light intensity for 500 ms using GFP filters (excitation 450–490 nm, emission 500–550 nm).

Topology Prediction—Topology predictions were conducted using four different programs: TMpred, TMHMM (21), HMMtop (22, 23), and TopPred2 (24).

Depletion of YidC and SecE—Strains carrying P_{BAD}-*yidC* or P_{BAD}-*secE* and an SPA- or 3× FLAG-tagged gene on the chromosome were grown overnight in LB containing 25 μg/ml of chloramphenicol, 30 μg/ml of kanamycin, and 0.2% arabinose at 37 °C. Overnight cultures were diluted 1:200 in fresh medium and grown for 2 h at 37 °C. Cells were collected by centrifugation and washed with 5 ml of LB lacking arabinose. After centrifugation to discard the wash, cells were resuspended and split into 2 flasks: one containing LB medium with 0.2% glucose and one with 0.2% arabinose. For supplemental Fig. S1, samples were taken at the indicated time points. For the subcellular fractionation, ~16 A_{600} units of cells were collected from each culture after 3 h of incubation at 37 °C.

RESULTS

Most Small Membrane Proteins Are Localized to the Inner Membrane—For several TM domain-containing small proteins it has been demonstrated that they fractionate with the pellet (membrane) fraction of *E. coli* lysates (2). The procedure used in these initial fractionations was not able to differentiate between a protein residing in the inner membrane of *E. coli* and one residing in the outer membrane. Because these small proteins are predicted to form α-helical TM domains instead of β-sheets, they were predicted to reside in the inner membrane. To determine the inner *versus* outer membrane localization for a set of the small proteins, we first employed a sucrose cushion fractionation protocol, which takes advantage of the fact that the inner membrane of *E. coli* has a lower density than the outer membrane so that the inner membrane will "float" on the cushion, whereas the outer membrane will sediment through the cushion (18).

As shown for seven control proteins tagged with the SPA-tag on the C terminus, the sucrose cushion protocol could be used to separate cell lysates into fractions that are enriched for soluble, inner membrane, and outer membrane proteins (Fig. 1, Table 1, and supplemental Table S4). Thus the soluble control proteins Pgm and YpfM, inner membrane control proteins AtpE, LepB, and Pf3, and outer membrane controls protein TolC and OmpA were most abundant in the expected fraction. There is some signal for each of the proteins in other fractions due to intrinsic limitations of subcellular fractionation (see "Experimental Procedures"). Nevertheless, the results with control proteins indicate that this protocol is able to show if a protein is localized predominantly in the cytosol, inner membrane, or outer membrane of *E. coli*.

We thus used the sucrose cushion protocol to examine the subcellular fractionation of SPA-tagged derivatives of two small proteins predicted to be soluble (YdfB, MntS), one predicted to be an amphipathic helix (AzuC), and 11 small proteins predicted to contain TM domains (Fig. 1, Table 1, and supplemental Table S4). As expected, the majority (61%) of YdfB-SPA is found in the soluble faction. In contrast, 43% of MntS-SPA was found in the soluble fraction and 55% was found in the inner membrane fraction suggesting that MntS may be loosely associated with the inner membrane or cycle on and off the mem-

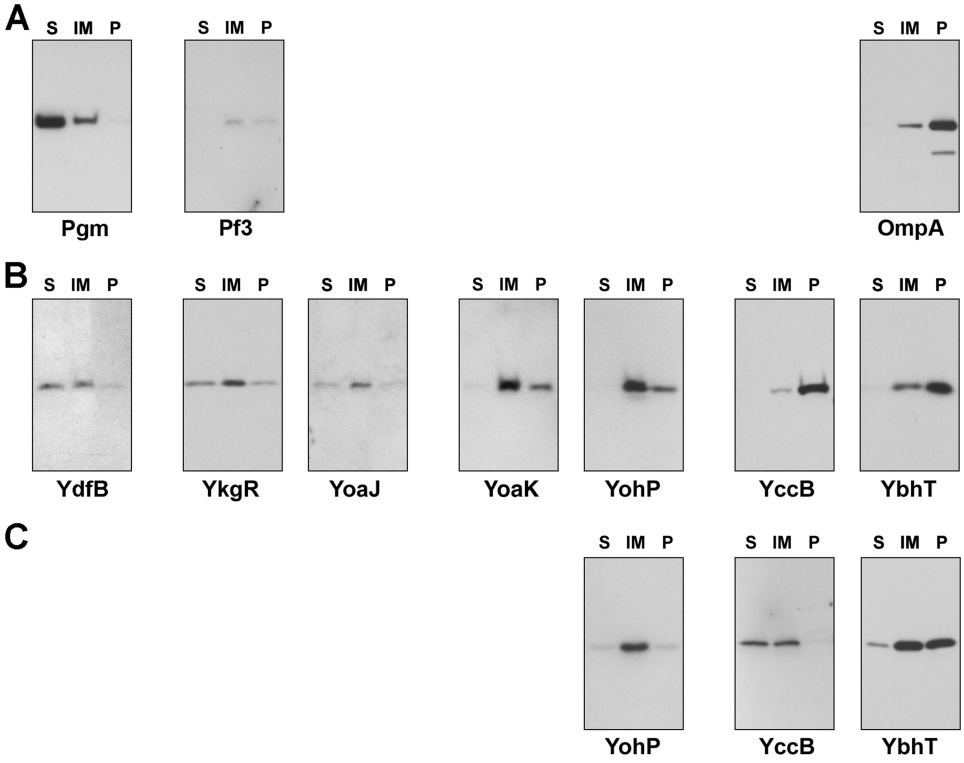

FIGURE 1. **Partitioning of small proteins between inner and outer membrane.** Examples of SPA-tagged proteins (control proteins, *A*; and small proteins, *B* and *C*) detected by immunoblot analysis after subcellular fractionation using a sucrose cushion (*A* and *B*) or Sarcosyl solubilization (*C*). The intensities of the bands from the soluble (*S*), inner membrane (*IM*), and pellet (*P*) fractions were measured for diluted samples and used to determine the relative abundance of a protein in each fraction. The results of these measurements are given under Table 1 and supplemental Tables S4 and S5.

TABLE 1

Partitioning of small proteins between inner and outer membrane fractions using a sucrose cushion

Protein	Expected location	*n*[a]	Soluble fraction[b]	Inner membrane fraction[b]	Outer membrane fraction (pellet)[b]
Pgm	Soluble	3	**67%**	30%	3%
YpfM	Soluble	3	**46%**	38%	17%
AtpE	Inner membrane	3	5%	**68%**	27%
LepB	Inner membrane	3	1%	**61%**	38%
Pf3	Inner membrane	3	4%	**52%**	44%
TolC	Outer membrane	3	15%	25%	**60%**
OmpA	Outer membrane	3	2%	20%	**78%**
YdfB	Soluble	3	**61%**	24%	15%
MntS	Soluble	3	43%	**55%**	2%
AzuC	Amphipathic	3	21%	**77%**	2%
YkgR	Inner membrane	3	27%	**58%**	15%
YoaJ	Inner membrane	3	17%	**57%**	26%
YnhF	Inner membrane	4	1%	**58%**	41%
YoaK	Inner membrane	3	5%	**61%**	34%
YohP	Inner membrane	3	5%	**63%**	32%
YncL	Inner membrane	3	21%	**67%**	12%
YpdK	Inner membrane	3	7%	**67%**	26%
YneM	Inner membrane	3	26%	**51%**	23%
YbgT	Inner membrane	5	2%	**53%**	45%
YccB	Inner membrane	3	2%	22%	**76%**
YbhT	Inner membrane	5	2%	42%	**56%**

[a] Number of repetitions. The numbers for individual experiments are given under supplemental Table S4.
[b] Percentage of small proteins in the different fractions were determined from immunoblot of the SPA-tagged proteins after subcellular fractionation. Bold values indicate the fraction in which the majority of the protein is found. Percentages represent the average of at least three independent experiments.

brane. AzuC-SPA was most abundant in the inner membrane fraction (77%) indicating this protein is bound to the cytoplasmic side of the inner membrane. Of the predicted TM proteins, nine (YkgR, YoaJ, YnhF, YoaK, YohP, YncL, YpdK, YneM, and YbgT) were found to fractionate with the inner membrane (51–77%); consistent with the prediction that TM domain-containing proteins would be found in the inner membrane. Unexpectedly, the signal for two of the proteins (76% of YccB-SPA and 56% of YbhT-SPA) was highest in the outer membrane fraction.

Using a Sarcosyl fractionation protocol that separates the inner and outer membranes based on differential solubilization, we also found that the YbgT-SPA and YohP-SPA proteins were predominantly (64 and 76%, respectively) in the inner membrane fraction along with YccB-SPA (51%) (Fig. 1*B* and supplemental Table S5). In contrast much of YbhT-SPA (43%) was in the outer membrane fraction. One possible explanation for this finding is that that YbhT is a component of the periplasm spanning AcrA-AcrB-TolC efflux pump.[5]

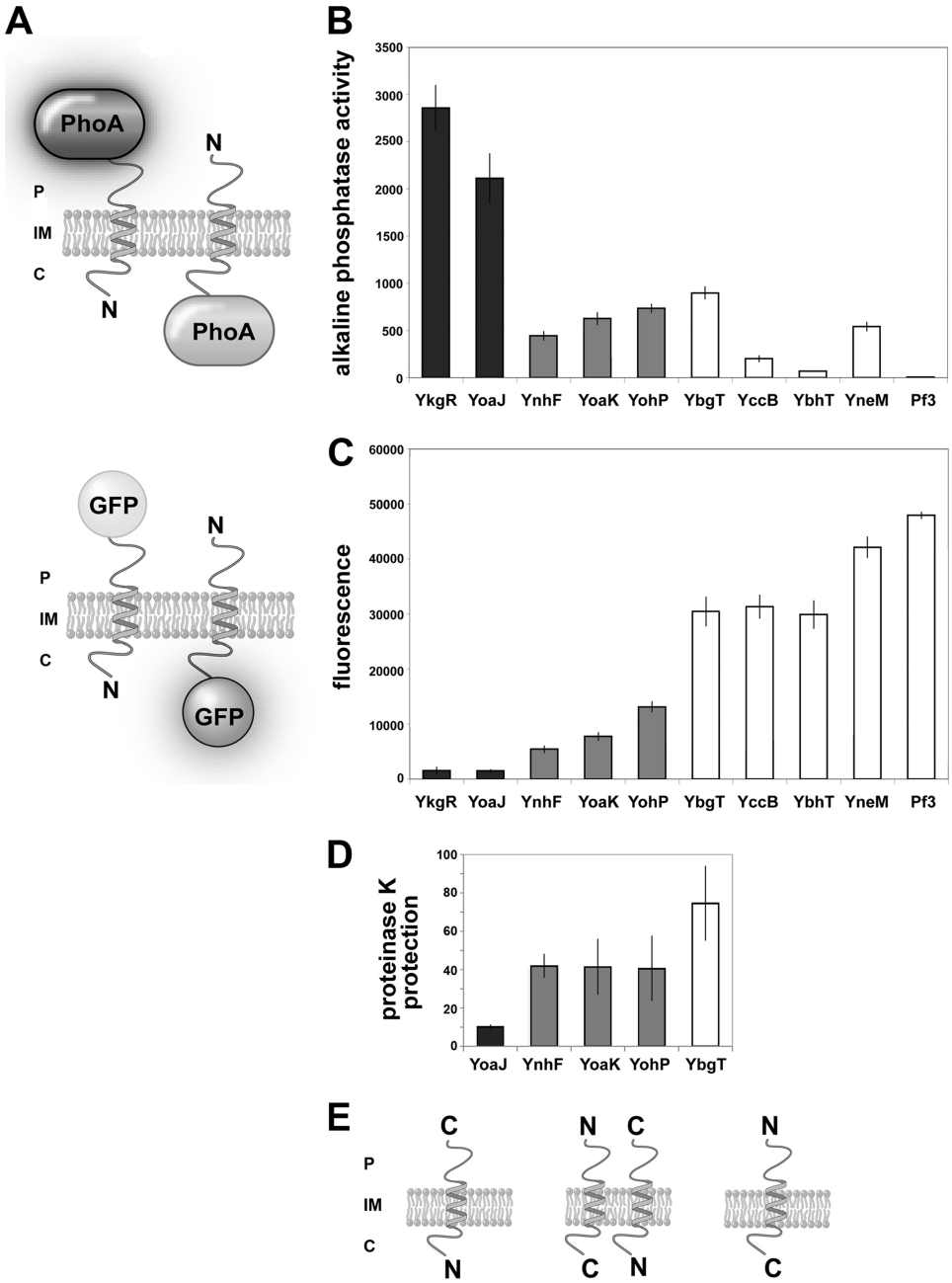

FIGURE 2. **Orientation of small proteins in the inner membrane.** *A,* orientation in which activity is observed for proteins C-terminal tagged with PhoA (C terminus in the periplasm) and GFP (C terminus in the cytoplasm). *B*, levels of alkaline phosphatase activity for all proteins tagged with PhoA on the C terminus. Activity obtained for the MG1655 Δ*phoA* parent strain carrying the empty plasmid pHA was subtracted from each value. The values shown are averages for the following number of independent samples: YkgR (9), YoaJ (9), YnhF (7), YoaK (8), YohP (6), YbgT (8), YccB (4), YbhT (8), YneM (6), and Pf3 (5). *C*, levels of fluorescence for all proteins tagged with GFP on the C terminus. Activity obtained for the BL21(DE3) pLysS parent strain carrying the empty plasmid pWALDO was subtracted from each value. The values shown are averages for the following number of independent samples: YkgR (6), YoaJ (10), YnhF (6), YoaK (7), YohP (10), YbgT (6), YccB (7), YbhT (7), YneM (9), and Pf3 (4). *D*, proteinase K accessibility of small proteins tagged with SPA on the C terminus. Percentage corresponds to the proportion of SPA protected from degradation. The values shown are averages for the following number of independent samples: YoaJ (3), YnhF (3), YoaK (3), YohP (4), and YbgT (5). For *B-D*, the *error bars* correspond to the standard deviation. *E,* suggested orientation of small proteins based on levels of alkaline phosphatase activity, GFP fluorescence, and proteinase K accessibility.

Consistent with this explanation, YbhT is localized to the inner membrane fraction in cells lacking AcrB (supplemental Table S5).

Small Membrane Proteins May Have Dual Orientations in the Membrane—The orientation of proteins in the membrane is important for their function because specific amino acids might be critical for an interaction with a partner molecule in a specific compartment such as the cytoplasm or periplasm. In a previous study, the topology of 600 inner membrane proteins in *E. coli* was examined using matched C-terminal fusion proteins to two different reporters of topology: the alkaline phosphatase (PhoA) and the green fluorescent protein (GFP) (12). PhoA is only active in the periplasm, whereas GFP is only fluorescent in the cytoplasm (Fig. 2*A*). Thus, a matched reporter pair of proteins with its C terminus in the periplasm will have high PhoA activity and a low GFP fluorescence. Conversely, a protein with

TABLE 2
Sequences of the small proteins with charged residues and topology predictions

Protein	Sequence[a]	Orientation[b]	TM pred[c]	TMHMM[c]	HMMTOP[c]	Top PRED2[c]
YoaJ	MKKTTIIMMGVAIIVVLGTELGWW	N_{in}	N_{in}	N_{in}	ND[d]	N_{in}
YkgR	MKENKVQQISHKLINIVVFVAIVEYAYLFLHFY	N_{in}	N_{out}	N_{in}	N_{in}	N_{in}
YnhF	MSTDLKFSLVTTIIVLGLIVAVGLTAALH	dual	N_{out}	N_{in}	N_{in}	N_{in}
YoaK	MRIGIIFPVVIFITAVVFLAWFFIGGYAAPGA	dual	N_{out}	N_{in}	N_{in}	N_{in}
YohP	MKIILWAVLIIFLIGLLVVTGVFKMIF	dual	N_{out}	N_{out}	N_{out}	unclear[e]
YbgT	MWYFAWILGTLLACSFGVITALALEHVESGKAGQEDI	N_{out}	N_{in}	N_{in}	N_{out}	N_{in}
YccB	MWYLLWFVGILLMCSLSTLVLVWLDPRLKS	N_{out}	N_{out}	N_{in}	N_{out}	N_{out}
YbhT	MLELLKSLVFAVIMVPVVMAIILGLIYGLGEVFNIFSGVGKKDQPGQNH	N_{out}	N_{in}	N_{in}	N_{in}	N_{out}
YneM	MLGNMNVFMAVLGIILFSGFLAAYFSHKWDD	N_{out}	N_{in}	N_{in}	N_{out}	unclear[e]
Pf3	MQSVITDVTGQLTAVQADITTIGGAIIVLAAVVLGIRWIKAQFF	N_{out}	N_{out}	N_{out}	N_{out}	N_{out}
YohP1	MKIILWAVLIIFLIGLLVVTGVFGMIF	N_{in}	N_{out}	N_{out}	N_{in}	N_{in}
YohP2	MGIILWAVLIIFLIGLLVVTGVFKMIF	N_{out}	N_{out}	N_{out}	N_{out}	N_{out}
YohP3	MGIILWAVLIIFLIGLLVVTGVFGMIF	dual	N_{out}	N_{out}	N_{out}	N_{in}

[a] Charged residues are in bold with positive and negative charges annotated in red and blue, respectively. The predicted transmembrane domains are highlighted in gray.
[b] Experimentally determined orientation using GFP and PhoA fusion proteins.
[c] Predictions of four different programs. For each program, the topology with the highest probability was reported in the table.
[d] Orientation could not be determined due to the short length of the protein.
[e] Prediction was ambigious.

its C terminus in the cytoplasm will have low PhoA activity and high GFP fluorescence. Daley *et al.* (12) focused on proteins longer than 100 residues containing at least two TM domains. Here we used the same strategy to analyze orientation of nine small proteins and the control protein Pf3, which has been reported to have an C_{in}-N_{out} orientation.

The small proteins we examined can be categorized in three classes: C_{out}-N_{in}, C_{in}-N_{out}, and possible dual orientation (Fig. 2, *B-E*). For YkgR and YoaJ, the PhoA fusions show high activity, whereas the corresponding GFP fusion show almost no activity indicating YkgR and YoaJ are oriented with the N terminus in the cytoplasm and the C terminus in the periplasm (C_{out}-N_{in}). In contrast, for YbgT, YccB, YbhT, and YneM, the PhoA fusions show low activity, whereas the corresponding GFP fusion show high activity similar to Pf3, consistent with an C_{in}-N_{out} orientation. For the third group of small proteins, YoaK, YohP, and YnhF, there were intermediate levels of both PhoA activity and GFP fluorescence. We interpret the intermediate levels to mean that YoaK, YohP, and YnhF can be inserted in both orientations (C_{out}-N_{in} and C_{in}-N_{out}). Six other proteins in *E. coli* have been shown to have dual orientations (12). All of these proteins are around 100 amino acids in length.

The differences in PhoA and GFP activity cannot solely be explained by variations in fusion protein expression because the PhoA fusions all have comparable levels of expression in immunoblot assays (supplemental Fig. S2). Similar immunoblot analysis showed that the YkgR-GFP, YnhF-GFP, and YbgT-GFP proteins are present at lower levels than the other GFP fusion proteins (supplemental Fig. S2), but the GFP fusions to YkgR, YbgT, and YnhF gave very different levels of fluorescence (Fig. 2*C*). Furthermore, localization of the GFP fusions in individual cells by fluorescence microscopy confirmed those fusions that give the highest fluorescence in the microplate assays were in the C_{in}-N_{out} orientation as expected (supplemental Fig. S3). Although it is possible that the larger PhoA and GFP tags can disturb the normal localization of the small proteins, the observation that the proteins show two different orientations indicates that the fusions do not force all of the small proteins into the same orientation.

We also tested the proteinase K sensitivity of YoaJ, YnhF, YoaK, YohP, and YbgT C-terminally tagged with the smaller SPA tag in spheroplasts of the corresponding strains. The results of these assays were consistent with the alkaline phosphatase and fluorescence assays; YoaJ-SPA showed high sensitivity (10% remaining) indicating accessibility of the C-terminal SPA tag on the exposed periplasmic side of the spheroplasts, YbgT-SPA had low sensitivity (75% remaining), whereas YnhF-SPA, YoaK-SPA, and YohP-SPA showed intermediate levels of sensitivity (41–42% remaining) (Fig. 2*D*). The tagged proteins all were similarly sensitive to digestion when the membranes were disrupted with Triton X-100.

Observed Orientation Has Varied Correlation with Predictions of Small Protein Orientation—The membrane localization of the small proteins was consistent with the fact that they were predicted to have a transmembrane domain by various programs (2). We used four programs (TM pred, TMHMM, HMMTOP 2.0, and TopPred 2 (26)) to evaluate if these programs could also accurately predict the topology of the small peptides in the membrane. However, the predictions for TM pred, TMHMM, HMMTOP 2.0, and TopPred 2, respectively, were consistent with only two, two, four, and five of the nine proteins assayed (Table 2). In many cases, the predicted orientation was opposite of the one observed. We also carried out predictions for the small protein-PhoA or small protein-GFP fusion proteins. The TMHMM and HMMTOP 2.0 programs predict a N_{in} orientation for all the fusion proteins, whereas TMHMM recognized the small ORFs as signal peptides. The general lack of correlation between the predictions and our

experimental results point to weaknesses in the programs for predicting the orientation of proteins containing only one TM domain.

Topology of Small Membrane Proteins Is Influenced by Charged Residues—The TopPred 2 program showed the highest success rate for predicting the orientation of the small proteins. This program is based on the "positive inside rule," which posits that positively charged residues such as lysine (Lys) and arginine (Arg) directly adjacent to the TM domain influences the orientation of the proteins in the membrane (24). All of the small proteins we examined have well conserved charged residues near the predicted TM helix (supplemental Fig. S4), and the orientations we observed are consistent with the positive inside rule. The C_{out}-N_{in} proteins have two (YoaJ) to three (YkgR) positive charges on the N terminus, whereas all of the C_{in}-N_{out} small proteins have positively charged residues on the C terminus. The exception is YbhT where positively charged residues were found on both sides, but for which there are more charged residues on the C-terminal end. For two small proteins with observed dual topology, the sequences show a weak asymmetry with respect to positive charges. YnhF and YoaK each have one positive charge on the N-terminal end. On the other hand, YohP has one positively charged residue (Lys) on either side of the TM domain, supporting the observation that YohP has a dual topology.

To test the influence of the lysine residues on the orientation of YohP, we replaced each with a glycine residue, either singly (YohP1 and YohP2) or in combination (YohP3) in the YohP-PhoA and YohP-GFP fusions (Fig. 3). The mutant YohP1, which has only one positive charge on the N-terminal side of the TM domain, has higher levels of PhoA activity and lower levels of GPF fluorescence and thus appears to adopt a C_{out}-N_{in} topology, as expected. In contrast, the mutant YohP2 fusion, which has a single positive charge on the C-terminal side of the TM domain, has lower levels of PhoA activity and higher levels of GPF fluorescence and appears to be inserted in a C_{in}-N_{out} orientation. When both lysines are mutated to glycine, the mutant YohP3 fusions show PhoA activity and GFP fluorescence comparable with the wild type YohP protein. The localization of the mutant YohP-GFP fusions in individual cells by microscopy is consistent with the assays with the YohP2-GFP fusion showing the clearest membrane localization (supplemental Fig. S5). Those results support the model that positively charged residues play a key role in the orientation of small proteins in the membrane. They also support the proposal that YohP has a dual topology in the inner membrane.

As an independent assay of the proposed dual topology of YohP and role of the two lysine residues, we examined proteinase K sensitivity of the PhoA tag for wild type and mutant derivatives of the YohP-PhoA fusions in spheroplasts (Fig. 3*D*). As expected, more than 99% of the YohP1-PhoA fusion protein was degraded because PhoA is proposed to be on the exposed periplasmic side of the membrane. For the YohP2-PhoA fusion protein, 61% remained after proteinase K treatment indicating that the majority of the PhoA is on the cytoplasmic side of the membrane and thus protected. The levels of YohP-PhoA fusion protein after digestion were intermediate (16% remaining) between the YohP1-PhoA and YohP2-PhoA proteins, again consistent with a dual orientation of the YohP-PhoA fusion. All of the fusions were equally sensitive to proteinase K treatment in the presence of Triton indicating that the differences in proteinase K sensitivity are due to differences in accessibility.

Small Protein Membrane Insertion Shows Differential Dependence on the Sec and YidC Translocation Systems—To determine which membrane insertion system is used by a particular small protein *in vivo*, we needed to disrupt each system and assay translocation of proteins by analyzing mutant lysate fractions. Both *secE* and *yidC* are essential genes in *E. coli* and cannot be deleted. Thus, we constructed strains where expression of the chromosomal *secE* and *yidC* genes were under control of an inducible P_{BAD} promoter. By repressing expression of the promoter in the presence of glucose and absence of arabinose, we were able to deplete the cells of these key translocation systems as reflected in the cessation of growth and rapid decrease in the levels of YidC and SecY, a component of the Sec complex shown to be reduced concomitantly with SecE (27). In each case the levels of the respective protein were reduced by more the 95% after 3 h, although SecE depletion also affected YidC levels slightly and YidC depletion also resulted in reduced SecY levels (supplemental Fig. S1). Subcellular fractionation based on membrane density, after 3 h of SecE or YidC depletion, showed that a Sec substrate (LepB-SPA) and a YidC substrate (AtpE-SPA) are present in higher abundance (>2-fold) in the soluble fraction of lysates of corresponding depleted cells indicating inhibited translocation to the membrane when the cells lack SecE or YidC (Fig. 4 and Table 3). Although membrane localization of AtpE-SPA was clearly affected by YidC, we did note that membrane localization was also decreased upon SecE depletion. We think this could be due to the somewhat reduced levels of YidC under the SecE depletion conditions, the different genetic background used in our studies, depletion of SecE rather than SecDF, or the fact that the proteins in our study were tagged and expressed from the chromosome (27, 28).

When the depletion and translocation assays were carried out for a selected set of small proteins, we observed a variety of results. Depletion of both SecE and YidC affected the translocation of three small proteins, YbhT-SPA, YbgT-SPA, and YoaJ-SPA, as reflected in the increased levels of protein in the soluble fraction (Fig. 4*A*, Table 3, and supplemental Table S6). In contrast, translocation of YoaK-SPA was primarily affected by YidC depletion, whereas translocation of YkgR-SPA was primarily affected by SecE depletion. The same dependence of YoaK on YidC also was observed when YidC was C-terminal tagged with the shorter 3× FLAG tag (Fig. 4*B*, Table 3, and supplemental Table S6). Intriguingly, the effect on translocation of YkgR-SPA upon SecE depletion was not characteristic of the effects of depletion on other small proteins. Specifically, there was not an increase in the soluble fraction of YkgR-SPA, but there was an increase in the outer membrane (pellet) fraction (Fig. 4*A*). This difference may be the result of YkgR-SPA forming aggregates that pellet through the sucrose cushion when YkgR-SPA translocation to the membrane is inhibited. Neither SecE nor YidC depletion affected the translocation of YohP-SPA. The results of the depletion assays dispute the prediction that YidC is solely required for the translocation of

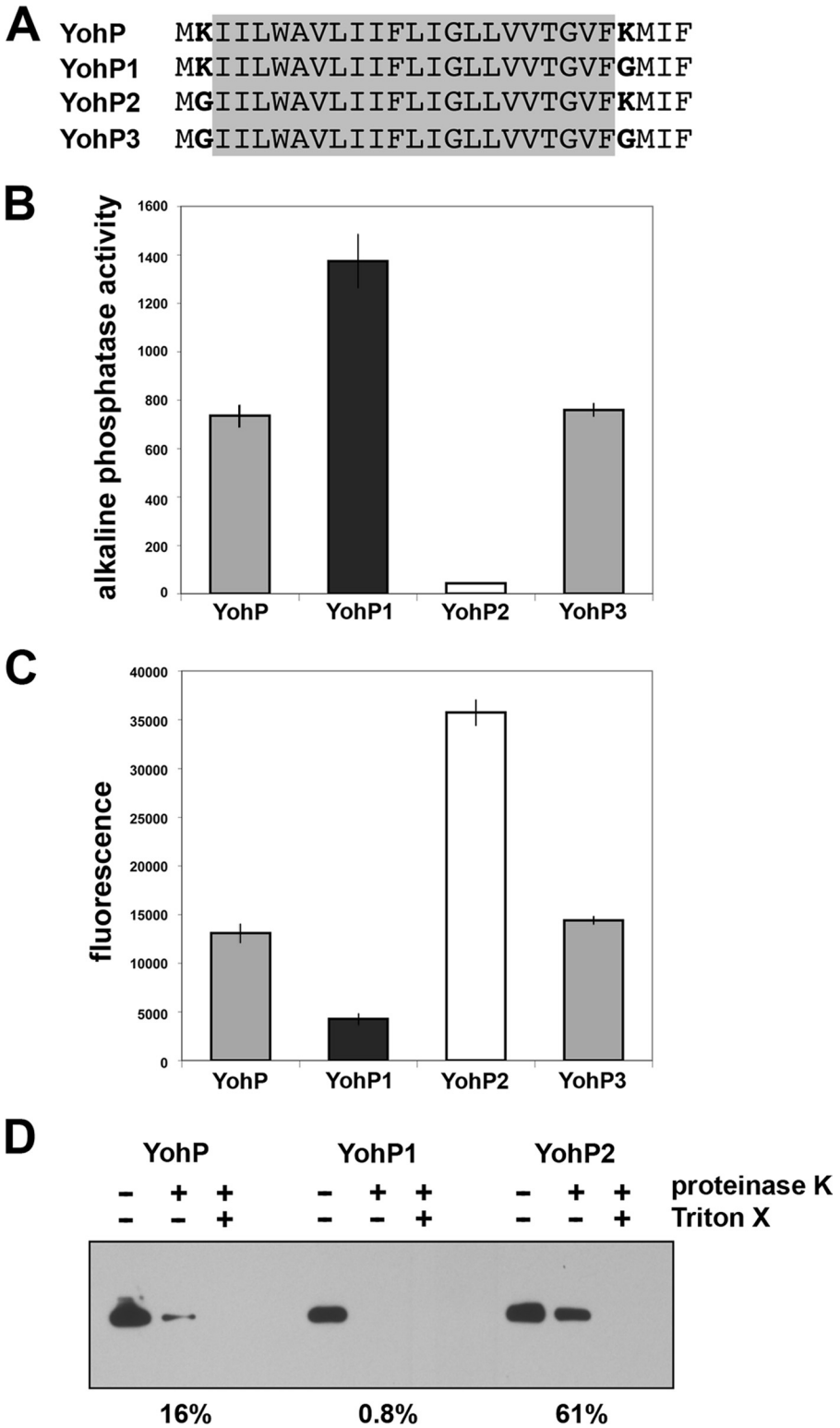

FIGURE 3. **Dual topology of the YohP protein.** *A,* sequences of wild type and mutant derivatives of YohP. Lysine residues are indicated in bold, and were replaced by a glycine: YohP1 ($Lys^{24} \rightarrow Gly$), YohP2 ($Lys^{2} \rightarrow Gly$), and YohP3 ($Lys^{2} \rightarrow Gly$ and $Lys^{24} \rightarrow Gly$). *B,* levels of alkaline phosphatase activity for wild type and mutant YohP proteins C-terminal tagged with PhoA. Activity obtained for the MG1655 Δ*phoA* parent strain carrying the empty plasmid pHA was subtracted from each value. The assays were carried out for the following number of independent samples: YohP (6), YohP1 (4), YohP2 (4), and YohP3 (4). *C,* levels of fluorescence for wild type and mutant YohP C-terminal tagged with GFP. Activity obtained for the BL21(DE3) pLysS parent strain carrying the empty plasmid pWALDO was subtracted from each value. The values shown are averages for the following number of independent samples: YohP (10), YohP1 (4), YohP2 (6), and YohP3 (3). The *error bars* correspond to the standard deviation. *D,* immunoblot analysis of PhoA fusions without and with proteinase K treatment. Spheroplasts of cells expressing the YohP-PhoA, YohP1-PhoA, and YohP2-PhoA fusion proteins were treated with proteinase K and Triton X-100 in the indicated samples. The percentage listed represents the proportion of PhoA protected from degradation.

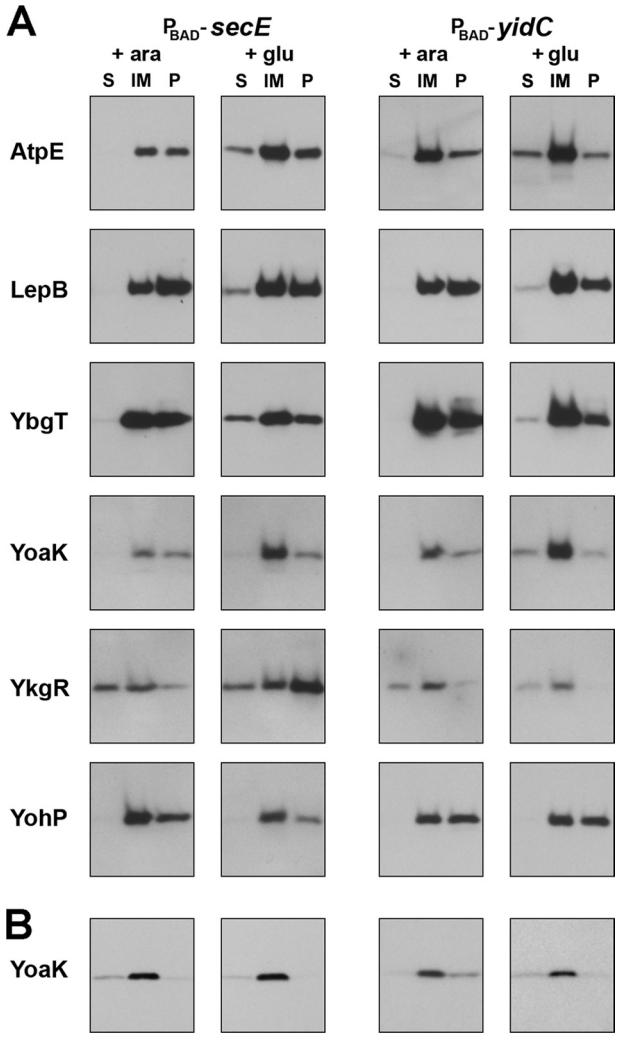

FIGURE 4. **Partitioning of proteins after depletion of YidC and SecE.** Examples of *A*, SPA-tagged proteins and *B*, 3× FLAG-tagged YoaK detected by immunoblot assays after subcellular fractionation using a sucrose cushion. Cells were grown in the presence of 0.2% arabinose to express SecE or YidC or in the presence of 0.2% glucose to deplete either YidC or SecE. The intensities of the bands from the soluble (*S*), inner membrane (*IM*), and pellet (*P*) fractions were measured for diluted samples and used to determine the relative abundance of the tagged protein in each fraction. For YkgR, ratio of the pellet fraction to all fractions combined (soluble, inner membrane, and pellet) was used to determine whether translocation was affected by depletion. For all other proteins, the ratio of the soluble fraction to all fractions combined was used to determine whether translocation was affected by depletion. The calculated ratios are given under Table 3 and supplemental Table S6.

these small proteins. Instead, the results suggest that TM domain-containing small proteins are recognized and translocated to the membrane using a variety of different mechanisms.

DISCUSSION

Although the expression and membrane localization of newly identified small TM domain-containing proteins in *E. coli* was confirmed (2), little was known about the functions and general properties of this class of proteins. The focus of this study was to investigate the physical properties of many of these proteins. Specifically, are the proteins located in the inner membrane or outer membrane, what is their orientation within the membrane, and how are they translocated to the membrane? Although the localization and orientation of a few small proteins such as MgrB (4) and MgtR (29) had been reported, no information about the translocation mechanism was available.

In *E. coli* and other Gram-negative bacteria, proteins that span the inner membrane typically have a hydrophobic α-helical domain, whereas the membrane spanning domains of outer membrane proteins are β-barrels comprised of β-sheets. About 65% of the recently discovered small proteins are predicted to contain a single α-helical TM domain and thus were hypothesized to be inner membrane proteins. Consistent with this hypothesis, the fractionation protocols used in this study showed that the majority of the proteins are most abundant in the inner membrane fraction. We suggest that the one exception, YbhT, might be retained in the outer membrane pellet fraction due to association with a large protein complex spanning the periplasm and inner and outer membranes.[3]

The differential results obtained for YccB with the sucrose cushion and Sarcosyl fractionation assays cannot be fully explained at this point. The observation that the YccB-GFP fusion shows high levels of fluorescence (Fig. 2*C*) is most consistent with inner membrane localization. The predominance of YccB-SPA in the outer membrane fraction in the density-based fractionation procedure might be due to association with a large protein complex as for YbhT.

Our results with C-terminal GFP, PhoA, and SPA fusions show that the orientation of these proteins within the inner membrane generally follows the "positive inside" rule. This rule is derived from the observation that most proteins have positively charged residues located on the cytoplasmic side of the membrane and that positive residues that are closer to the TM domain have a stronger effect on the orientation of the protein (24). We propose that expanded experimental determination of small protein orientation can be used to improve membrane protein topology prediction programs.

At least one of the small proteins, YohP, has a dual orientation. Thus far, very few proteins have been reported to have a dual topology, but all of them are small (around 120 amino acids) (12). The dual topology protein that has been studied most extensively is EmrE, a small multidrug transporter. In the EmrE example, the cohabitation of the two topologies has been shown to be essential for the function of the protein (30). Whether the dual topology is crucial to YohP function remains to be determined.

In *E. coli*, two membrane insertion systems for membrane proteins have been characterized: the Sec translocation system and the YidC translocase. Most proteins in *E. coli* utilize the Sec system to either insert into or translocate across the membrane via the SecYEG pore. Accessory proteins aid in various steps in the insertion and translocation processes and, in the case of YidC, the release of the protein substrate into the membrane. In addition to its accessory role in the Sec system, YidC has been shown to act as a membrane insertase on its own (9). Substrates reported to be dependent only on YidC *in vivo* (9, 10, 27, 28, 31) tend to consist of only one or two transmembrane domains. We thus anticipated that the small proteins characterized in this study would utilize this system, but noted that the two YidC-only dependent substrates we examined, Pf3-SPA and AtpE-SPA, were affected by both YidC and SecE depletion in our experiments.

TABLE 3
Effect of SecE and YidC depletion on protein translocation

Protein	Depletion	Ratio[a]	Standard deviation	*n*[b]	*Z*-value[c]	Confidence <2[c]	Confidence >2[c]
LepB-SPA	SecE	5.7	1.8	7	3.6	0%	**100%**
LepB-SPA	YidC	5.6	2.4	3	3.6	0%	**100%**
AtpE-SPA	SecE	7.3	1.6	3	5.8	0%	**100%**
AtpE-SPA	YidC	3.3	1.1	8	3.1	0%	**100%**
YbhT-SPA	SecE	5.6	3.2	5	2.5	2%	**98%**
YbhT-SPA	YidC	6.7	2.7	5	3.9	0%	**100%**
YbgT-SPA	SecE	6.9	1.7	3	4.9	0%	**100%**
YbgT-SPA	YidC	6.3	3.3	3	2.2	1%	**99%**
YoaJ-SPA	SecE	3.6	1.3	4	2.6	1%	**99%**
YoaJ-SPA	YidC	2.8	0.74	5	2.5	2%	**98%**
YoaK-SPA	SecE	1.5	0.55	6	−2.2	**98%**	2%
YoaK-SPA	YidC	5.6	1.5	4	4.7	0%	**100%**
YkgR-SPA	SecE	3.6	1.3	3	2.1	2%[d]	**98%**[d]
YkgR-SPA	YidC	0.61	0.59	4	−4.7	**100%**[d]	0%[d]
YohP-SPA	SecE	1.1	0.24	4	−7.5	**100%**	0%
YohP-SPA	YidC	1.4	0.29	5	−5.0	**100%**	0%
YoaK-3× FLAG	SecE	1.3	0.59	3	−2.2	**99%**	1%
YoaK-3× FLAG	YidC	3.3	1.2	4	2.1	2%	**98%**

[a] Ratio of signal in soluble fraction divided by signal in all fractions (soluble, inner membrane fraction, and pellet) for wild type compared to the depleted strain. A ratio of "1" indicates that there is no effect of YidC or SecE depletion on the ability of the SPA- and 3×FLAG-tagged protein to be inserted into the membrane. A value of "2" or greater corresponds to more protein in the soluble fraction in the depleted strain indicating that the translocation of the protein to the membrane is impaired.
[b] Number of repetitions. The numbers for individual experiments are given under supplemental Table S6.
[c] Data were analyzed assuming a normal distribution. The *Z*-value was determined using the distance of the calculated average ratio from a value of 2 divided by the standard error. The *Z*-value was then used to determine the confidence that the actual depleted *versus* wild type ratio is either less than 2 or greater than 2. Each protein/depletion system combination produced results where the % confidence was greater than 97% that the true value was either less than 2 or greater than 2.
[d] For those samples, the ratio of the signal in the pellet fraction was divided by the signal in all fractions (soluble, inner membrane fraction and pellet) for wild type compared to the depleted strain. The signal in the soluble fraction did not change.

Contrary to our expectation, our SecE and YidC depletion assays indicate that the small proteins targeted in this study appear to be inserted into the inner membrane via a variety of mechanisms. The subcellular localization of three proteins (YbhT-SPA, YbgT-SPA, and YoaJ-SPA) was affected by both SecE and YidC depletion. We suggest that these proteins utilize the Sec translocation system and that the YidC may impact translocation due to its role as an accessory protein of the Sec system. In contrast, YkgR-SPA localization was primarily affected by SecE depletion. This suggests that YkgR is also a Sec substrate, but might not require YidC as an accessory protein, although the reason for this is unclear. YoaK-SPA was primarily affected by YidC depletion, indicating its independence from the Sec system and suggesting that YidC on its own might mediate membrane insertion for YoaK. Neither SecE nor YidC depletion had an effect on YohP-SPA indicating that YohP is not a substrate for either of these translocation systems.

Possible explanations for observation that YohP translocation was not affected by either SecE or YidC depletion are that YohP is able to spontaneously insert itself into the membrane, can be inserted by either system indiscriminately, or requires another translocation system in *E. coli*. One other characterized translocation system is the twin-arginine translocation (Tat) pathway. The Tat system is unique from the SecE- and YidC-dependent systems in that it translocates fully folded proteins across the inner membrane (32, 33). We considered it unlikely that the small proteins in this study would utilize the Tat system as none of them contain the prerequisite twin-arginine motif. However, it has been shown that if a protein is present in a complex that is translocated by the Tat system, then the rest of the complex can "piggyback" across the membrane (34). The same could be true for YohP; the small protein could conceivably be translocated across the membrane together with a Tat substrate. It is also possible that another translocation system remains to be discovered. The KdpD sensor kinase also has been reported to insert in the inner membrane independently of the Sec translocase and YidC (35). Similarly, one group has found that membrane insertion of the MscL does not require Sec or YidC, although in this case, YidC appears to assist the assembly of the homopentameric MscL complex within the membrane (36).

On the whole, these results indicate that the prediction that all small TM domain-containing proteins would utilize the YidC insertase due to their size might not be correct. A caveat to these results is that the small proteins tested in our depletion assays are tagged with an SPA tag. Although it is possible that the SPA tag could be affecting small protein translocation, particularly when the tag is in the periplasm, we suggest the tag is not having a significant effect. Previous work on translocation has shown that the information for determining the translocation mechanism most often is located toward the N terminus of the protein and our SPA fusions are on the C terminus. In addition, other genetic assays have shown that the SPA tag does not interfere with YbhT function and neither the SPA nor the PhoA tag affects YneM function.[5,7] In addition, YoaK dependence on YidC was also observed with the shorter and less charged 3× FLAG. Improvements in small protein detection, along with *in vivo* pulse-chase and *in vitro* insertion assays should help to further elucidate the mechanisms of small protein insertion into membranes.

The small sizes of the proteins characterized in our studies bring up other general questions such as how the proteins, some of which are shorter than the ribosome exit channel, leave the ribosome and whether they are bound by protein chaperones, some of which commonly deliver nascent peptides to the translocation machineries. Membrane insertion by the Sec system usually requires that the nascent protein be targeted to the

[7] F. Fontaine, unpublished results.

membrane by the signal recognition particle and its receptor FtsY (8). Membrane insertion by the YidC-dependent substrates M13 and Pf3 reportedly does not require the signal recognition particle targeting (37, 38) although insertion of the Sec- and YidC-independent substrate MscL does (36). We expect there to be similar diversity in the need for protein chaperones and signal recognition particles among the small membrane proteins. Further studies of small membrane protein biogenesis should give more general insights into protein release from the ribosome, folding within the cell, and membrane insertion as well as provide interesting clues to the functions of the small proteins.

Acknowledgments—We thank H. Bernstein for the YidC and SecY antiserum, D. Daley for plasmids pHA4 and pWALDO, and H. Bernstein, R. Hegde, M. Hemm, T. Hessa, and K. Ramamurthi for valuable experimental advice.

REFERENCES

1. Garbis, S., Lubec, G., and Fountoulakis, M. (2005) *J. Chromatogr. A* **1077,** 1–18
2. Hemm, M. R., Paul, B. J., Schneider, T. D., Storz, G., and Rudd, K. E. (2008) *Mol. Microbiol.* **70,** 1487–1501
3. Hobbs, E. C., Fontaine, F., Yin, X., and Storz, G. (2011) *Curr. Opin. Microbiol.* **14,** 167–173
4. Lippa, A. M., and Goulian, M. (2009) *PLoS Genet.* **5,** e1000788
5. Rowland, S. L., Burkholder, W. F., Cunningham, K. A., Maciejewski, M. W., Grossman, A. D., and King, G. F. (2004) *Mol. Cell* **13,** 689–701
6. Ramamurthi, K. S., Lecuyer, S., Stone, H. A., and Losick, R. (2009) *Science* **323,** 1354–1357
7. Gassel, M., Möllenkamp, T., Puppe, W., and Altendorf, K. (1999) *J. Biol. Chem.* **274,** 37901–37907
8. Driessen, A. J., and Nouwen, N. (2008) *Annu. Rev. Biochem.* **77,** 643–667
9. Samuelson, J. C., Chen, M., Jiang, F., Möller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) *Nature* **406,** 637–641
10. Yi, L., Jiang, F., Chen, M., Cain, B., Bolhuis, A., and Dalbey, R. E. (2003) *Biochemistry* **42,** 10537–10544
11. van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A. J. (2004) *J. Cell Biol.* **165,** 213–222
12. Daley, D. O., Rapp, M., Granseth, E., Melén, K., Drew, D., and von Heijne, G. (2005) *Science* **308,** 1321–1333
13. Rapp, M., Drew, D., Daley, D. O., Nilsson, J., Carvalho, T., Melén, K., De Gier, J. W., and Von Heijne, G. (2004) *Protein Sci.* **13,** 937–945
14. Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97,** 5978–5983
15. Morita, T., Kawamoto, H., Mizota, T., Inada, T., and Aiba, H. (2004) *Mol. Microbiol.* **54,** 1063–1075
16. Khlebnikov, A., Datsenko, K. A., Skaug, T., Wanner, B. L., and Keasling, J. D. (2001) *Microbiology* **147,** 3241–3247
17. Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., and Oppenheim, A. (2007) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) Vol. 79, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ
18. Rhoads, D. B., Tai, P. C., and Davis, B. D. (1984) *J. Bacteriol.* **159,** 63–70
19. Qi, H. Y., and Bernstein, H. D. (1999) *J. Biol. Chem.* **274,** 8993–8997
20. Manoil, C. (1991) *Methods Cell Biol.* **34,** 61–75
21. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) *Proc. Intl. Conf. Intell. Syst. Mol. Biol.* **6,** 175–182
22. Tusnády, G. E., and Simon, I. (1998) *J. Mol. Biol.* **283,** 489–506
23. Tusnády, G. E., and Simon, I. (2001) *Bioinformatics* **17,** 849–850
24. von Heijne, G. (1992) *J. Mol. Biol.* **225,** 487–494
25. Xie, K., and Dalbey, R. E. (2008) *Nat. Rev. Microbiol.* **6,** 234–244
26. Ikeda, M., Arai, M., Lao, D. M., and Shimizu, T. (2002) *In Silico Biol.* **2,** 19–33
27. Baars, L., Wagner, S., Wickström, D., Klepsch, M., Ytterberg, A. J., van Wijk, K. J., and de Gier, J. W. (2008) *J. Bacteriol.* **190,** 3505–3525
28. Yi, L., Celebi, N., Chen, M., and Dalbey, R. E. (2004) *J. Biol. Chem.* **279,** 39260–39267
29. Alix, E., and Blanc-Potard, A. B. (2008) *EMBO J.* **27,** 546–557
30. Rapp, M., Seppälä, S., Granseth, E., and von Heijne, G. (2007) *Science* **315,** 1282–1284
31. Chen, M., Xie, K., Nouwen, N., Driessen, A. J., and Dalbey, R. E. (2003) *J. Biol. Chem.* **278,** 23295–23300
32. Sargent, F. (2007) *Biochem. Soc. Trans.* **35,** 835–847
33. Hatzixanthis, K., Palmer, T., and Sargent, F. (2003) *Mol. Microbiol.* **49,** 1377–1390
34. Rodrigue, A., Chanal, A., Beck, K., Müller, M., and Wu, L. F. (1999) *J. Biol. Chem.* **274,** 13223–13228
35. Facey, S. J., and Kuhn, A. (2003) *Eur. J. Biochem.* **270,** 1724–1734
36. Pop, O. I., Soprova, Z., Koningstein, G., Scheffers, D. J., van Ulsen, P., Wickström, D., de Gier, J. W., and Luirink, J. (2009) *FEBS J.* **276,** 4891–4899
37. Chen, M., Samuelson, J. C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R. E. (2002) *J. Biol. Chem.* **277,** 7670–7675
38. de Gier, J. W., Scotti, P. A., Sääf, A., Valent, Q. A., Kuhn, A., Luirink, J., and von Heijne, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95,** 14646–14651